Massachusetts Institute of Technology
Department of Economics
Working Paper Series

CAPITOL INCOME TAXES WITH HETEROGENOUS DISCOUNT RATES

Peter A. Diamond
Johannes Spinnewijn

Working Paper 09-22
July 14, 2009

Room E52-251
50 Memorial Drive
Cambridge, MA 02142

Capital Income Taxes with Heterogeneous Discount Rates*

Peter Diamond† Johannes Spinnewijn‡
MIT MIT

July 14, 2009

Abstract

With heterogeneity in both skills and preferences for the future, the Atkinson-Stiglitz result that savings should not be taxed with optimal taxation of earnings does not hold. Empirical evidence shows that on average people with higher skills save at higher rates. Saez (2002) suggests that with such positive correlation taxing savings can increase welfare. This paper analyzes this issue in a model with less than perfect correlation between ability and preference for the future. To have multiple types at the same earnings level, the number of types of jobs in the economy is restricted. Key to the analysis is that types who value future consumption less are more tempted to switch to a lower earning job. We show that introducing both a small savings tax on the high earners and a small savings subsidy on the low earners increase welfare, regardless of the correlation between ability and preferences for the future. However, a uniform savings tax, as in the Nordic dual income tax, increases welfare only if that correlation is sufficiently high. There are also some results on optimal taxes that parallel the results on introducing small taxes.

Keywords: Optimal Taxation, Capital Income, Discount Rates
JEL Classification Number: H21

*We are grateful to Jesse Edgerton, Louis Kaplow, Emmanuel Saez, Ivan Werning and seminar participants at MIT, Berkeley, Stockholm University, the Stockholm School of Economics and the NBER for valuable comments, and the National Science Foundation for financial support (Award 0648741). The research reported herein was supported by the Center for Retirement Research at Boston College pursuant to a grant from the U.S. Social Security Administration funded as part of the Retirement Research Consortium. The findings and conclusions are solely those of the authors and should not be construed as representing the opinions or policy of the Social Security Administration or any agency of the Federal Government; or the Center for Retirement Research at Boston College.

†Department of Economics, E52-344, 50 Memorial Drive, Cambridge, MA 02142, pdiamond@mit.edu.
‡Department of Economics, E51-090, 50 Memorial Drive, Cambridge, MA 02142, jspinnew@mit.edu.

1 Introduction

The Atkinson-Stiglitz (1976) theorem shows that when the available tax tools include non-linear earnings taxes, optimal taxation is inconsistent with taxing savings when two key assumptions are satisfied: (1) that all consumers have preferences that are separable between consumption and labor and (2) that all consumers have the same sub-utility function of consumption. Empirical evidence suggests that on average those with higher skills save at higher rates (Dynan, Skinner and Zeldes, 2004, Banks and Diamond, 2008). We therefore relax the second condition and analyze the taxation of savings with heterogeneity in both skill and savings propensity. We consider both uniform and earnings-varying taxation of savings.

This paper uses a simple model in which the number of types of jobs in the economy is restricted. This sheds light on the desirability of earnings-dependent savings taxes and the role of the positive correlation between skill and savings propensity. The paper provides an argument for making the taxation of savings progressive in earnings. In a two-skills model, we find that the savings of the high earners should be taxed, whereas the savings of the low earners should be subsidized. This result is independent of the correlation between ability and discount factors, provided that the optimum has the high skilled workers on the more productive job. A uniform savings tax, however, only increases welfare if that correlation is sufficiently high.

Our paper builds on the analysis in Saez (2002). He derives conditions on endogenous variables to sign the effect on social welfare of introducing a uniform commodity tax or a subsidy, when consumers have heterogeneous sub-utility functions of consumption. With an optimal non-linear earnings tax, a small tax on savings increases welfare if either the net marginal social value is negatively correlated with savings, conditional on earnings, or on average those who choose to earn less save less than those who choose to earn more, if restricted to the same earnings. By restricting the number of types of jobs, we analyze the importance of the (exogenous) correlation between skills and savings preferences for the taxation of savings.

Primary attention is focused on a model with four worker types - with two discount factors and two skill levels. Thus we are examining a particular example of a multidimensional screening problem. The model assumes the existence of two jobs, rather than the standard model where each worker can select the number of hours to be worked.[1] This results in a setting where workers with the same skill but different discount factors choose the same job and so have the same earnings. With the introduction of earnings-related savings tax rates,

[1] A limited number of jobs was assumed in Diamond (2006).

they are subject to the same tax rates. We assume that at the optimum both high-skill types work at the high-skill job and that redistribution from high earners to low earners is the important redistribution. Given these assumptions social welfare increases with the introduction of a tax on the savings of high earners and with the introduction of a subsidy on the savings of low earners. The relative frequencies of the four types in the population plays no role in the derivation of this result, conditional on the assumed structure of the optimum.

The underlying assumption is that those valuing the future more are more willing to work than those valuing the future less, conditional on the disutility of work. This means that an incentive compatibility (IC) constraint just binding on a high skill worker with low value for the future is not binding on a high skill worker with high value for the future. Earnings-dependent taxes and subsidies on savings allow an increase in redistribution by targeting types in a given job with saving preferences different than those of types who are just tempted to switch jobs. In particular, introducing taxation of savings of high earners (and transferring the revenue back equally to all high earners) eases the binding IC constraint since it transfers resources from the high saver to the low saver for whom the IC constraint is binding. Introducing a subsidy on savings for low earners (financed by equal taxation on all low earners) also eases the binding IC constraint by making switching to the lower job less attractive to the high earner with low savings. In extensions, the case for taxing the savings of high earners appears to be more robust than the case for subsidizing the savings of low earners. While the focus of the paper is the introduction of small taxes, we also consider optimal taxes under stronger assumptions.[2]

The assumption that those with less discounting of the future are more willing to work is in line with standard modeling, representing preferences by $u(x) + \delta_i u(c) - v(z/n_i)$. An alternative specification $\frac{1}{\delta_i} u(x) + u(c) - v(z/n_i)$ would imply the exact opposite. That is, types with higher δ_i prefer to save more, but to work less. We examine some empirical support for our assumption, using data from the Survey of Consumer Finances (SCF). We find that conditional on education and age, people with higher discount factors tend to earn more. To proxy for the discount factor, we use reported savings and the time horizon people report having in mind when making spending and savings decision. We also use these proxies to revisit the positive correlation between skills and savings propensities.

This paper contributes to the literature on the optimal choice of the tax base and the joint taxation of labor and capital incomes in particular. Banks and Diamond (2008) review the literature on the inclusion of capital income in the tax base. Gordon (2004) and Gordon

[2]The analysis assumes rational savings by all workers. Concern about too little individual savings is also relevant for retirement savings policies.

and Kopczuk (2008) argue that capital income reveals information about earnings ability and thus should be included in the tax base. Blomquist and Christiansen (2008) analyze how people with different skills and different preferences for leisure who cannot be separated with an income tax, may be separated with a commodity tax. The four-types model with hours chosen by workers has been studied by Tenhunen and Tuomala (2008), which calculates a set of examples, but explores the analytics only in two- and three-type models. They consider both welfarist and paternalist objective functions. We relate the results in their calculated examples to some of our results below. We focus on the four-types model since the result in a two-types model, while striking, does not seem relevant for policy inferences.[3] While the focus of this paper is on capital taxation, the intuition generalizes to the taxation of other commodities for which the preferences are heterogeneous, since this heterogeneity may impact the labor choice as well (Kaplow, 2008a).

The paper is organized as follows. Section 2 sets up the model with four types and two jobs. Section 3 characterizes respectively the first best and the restricted first best, referring to no taxation of savings and an 'equal job, equal pay' restriction. Section 4 introduces incentive compatibility constraints and characterizes the second best including the introduction of earnings-varying savings tax rates. Optimal savings tax rates are also considered. Section 5 considers a uniform savings tax, rather than one varying with the level of earnings. For comparison, Section 6 reviews a two-types model. Section 7 discusses empirical support for the assumptions and Section 8 has concluding remarks.

2 Model

We consider a model with two periods. Agents consume in both periods, but work only in the first period. Preferences are assumed to be separable over time and between consumption and work. Denoting first period consumption by x, second period consumption by c, and earnings by z, preferences satisfy

$$U(x, c, z) = u(x) + \delta u(c) - v(z/n),$$

with $u' > 0, u'' < 0$ and $v' > 0, v'' > 0$. An agent's ability n determines the disutility of producing output z. An agent's preference for future consumption depends on the discount factor δ.

We consider heterogeneity in both ability n and preference for future consumption δ.

[3]Kocherlakota (2005) provides an argument for regressive earnings-varying wealth taxation. He analyzes a model with asymmetric information about stochastically evolving skills, which is not present in this model.

Although robust insights for optimal taxation have been derived in models with two types, considering heterogeneity in two parameters in a model with two types implies perfect correlation between the two parameters. The inference based on a simple two-types economy, although simple, may therefore be misleading. In order to allow for imperfect correlation, we consider a four-types model. We denote the four types by ll, lh, hl, hh with frequencies f_{ij} and welfare-weights η_{ij}. The first two types have low ability n_l, but differ in discount factors δ_l and δ_h, with $\delta_h > \delta_l$. The second two types have high ability n_h, with $n_h > n_l$, and also differ in discount factors δ_l and δ_h.

	high discount factor δ_h	low discount factor δ_l
high ability n_h	**hh**	**hl**
low ability n_l	**lh**	**ll**

There are only two jobs in the economy, h and l. The output from a job is independent of the worker's type, while the disutility of holding a job varies with ability. The low-ability types can only hold the low job. The high-ability types can hold either job. We assume that redistribution to the low-skilled types is sufficiently important and the type mix sufficiently balanced that all high-skilled workers hold high-skilled jobs at the various optima analyzed. This requires a restriction on the weights in the social welfare functions and the population distribution, which we do not explore.

We begin with the first best, which differs from the usual treatment in that the output produced on a job is the same for everyone holding the job. We assume a linear technology. The first best has the property that there is no marginal taxation of savings. Then we consider a *restricted first best* (the term 'first best' refers to a lack of incentive compatibility constraints, the term 'restricted' means limited tax tools, but not limited by IC constraints) with zero taxation of savings and the requirement that everyone holding a job receives the same pay (no taxes based on identity, only on potential earnings). We calculate whether social welfare can be improved by taxing or subsidizing savings.

We then turn to the second best, with taxes based on earnings, not potential earnings, so that there is an incentive compatibility constraint. We assume a zero taxation of savings restriction, thus preserving the condition of equal pay for equal work. Again we ask about potential gains from taxing or subsidizing savings.

3 First Best

In the first best, each worker is assigned to the matching job and the social welfare function is maximized with respect to the type-specific consumption levels in the first and second periods and the job-specific output levels, subject to a resource constraint. With the welfare weight of type ij of η_{ij}, the first best solves:

$$\begin{aligned} &\text{Maximize}_{x,c,z} \quad \sum f_{ij}\eta_{ij}\left(u\left[x_{ij}\right]+\delta_j u\left[c_{ij}\right]-v\left[z_i/n_i\right]\right) \\ &\text{subject to:} \qquad E+\sum f_{ij}\left(x_{ij}+R^{-1}c_{ij}-z_i\right)\leq 0 \end{aligned} \tag{1}$$

Forming a Lagrangian with λ the Lagrange multiplier for the resource constraint, we have

$$\mathcal{L}=\sum_{i,j} f_{ij}\eta_{ij}\left(u\left[x_{ij}\right]+\delta_j u\left[c_{ij}\right]-v\left[z_i/n_i\right]\right)-\lambda\sum_{i,j} f_{ij}\left(x_{ij}+R^{-1}c_{ij}-z_i\right)$$

We define the net marginal social value of first period consumption for an individual of type ij as

$$g_{ij}\equiv\eta_{ij}u'\left[x_{ij}\right]-\lambda.$$

Along the relevant portion of the social welfare optima, we have the following properties:

$$g_{ij}=0 \text{ and } u'\left[x_{ij}\right]=\delta_j Ru'\left[c_{ij}\right] \text{ for all } i,j, \text{ and}$$

$$\left(f_{il}\eta_{il}+f_{ih}\eta_{ih}\right)\frac{v'\left[z_i/n_i\right]}{n_i}=\left(f_{il}+f_{ih}\right)\lambda=f_{il}\eta_{il}u'\left[x_{il}\right]+f_{ih}\eta_{ih}u'\left[x_{ih}\right],$$

for both the high-skilled and the low-skilled jobs. The net marginal social value of first period consumption for each type equals 0 and the saving of each type is undistorted. Given that the required output for a given job is independent of an individual's type, the earnings are marginally distorted upward for one discount-factor type and downward for the other type, since $u'\left[x_{il}\right]\neq u'\left[x_{ih}\right]$, unless the welfare weights satisfy $\eta_{il}=\eta_{ih}$. The output is undistorted 'on average' though.

3.1 Restricted First Best: Equal Pay for Equal Work and No Taxation of Savings

If the (after-tax) earnings on a job, y_i, is restricted to be type-independent and savings can not be taxed, there are further constraints, which we approach using the indirect utility-of-

consumption function, $w_j[y, R]$. This function satisfies

$$w_j[y, R] \equiv \max u[x] + \delta_j u[c]$$

$$\text{subject to: } \quad x + R^{-1}c = y.$$

For later use, we note that

$$\begin{aligned} \frac{\partial w_j}{\partial y} &= u'[x] \\ \frac{\partial w_j}{\partial R} &= R^{-2}cu'[x] = R^{-1}(y-x)u'[x] = R^{-1}s_j[y, R]u'[x] \end{aligned}$$

where $s_j[y, R]$ is the savings function of someone with discount factor δ_j.

We continue to assume that the welfare weights and population fractions are such that all high skilled are on the more productive job at the optimum. The restricted first best solves the following problem,

$$\begin{aligned} &\text{Maximize}_{y,z} \quad \sum f_{ij}\eta_{ij}(w_j[y_i, R] - v[z_i/n_i]) \\ &\text{subject to:} \qquad E + \sum f_{ij}(y_i - z_i) \leq 0 \end{aligned} \tag{2}$$

Forming a Lagrangian, we have

$$\mathcal{L} = \sum_{i,j} f_{ij}\eta_{ij}(w_j[y_i, R] - v[z_i/n_i]) - \lambda \sum_{i,j} f_{ij}(y_i - z_i).$$

The first order conditions (FOC) are

$$\begin{aligned} \sum_j f_{ij}\eta_{ij}u'[x_{ij}] &= \lambda \sum_j f_{ij} \\ \sum_j f_{ij}\eta_{ij}v'[z_i/n_i]/n_i &= \lambda \sum_j f_{ij}, \end{aligned}$$

for $i = h, l$. Recalling the definition of the net marginal social utility, $g_{ij} \equiv \eta_{ij}u'[x_{ij}] - \lambda$, the population-weighted values add to zero at each job,

$$\sum_j f_{ij}g_{ij} = 0 \text{ for } i = h, l.$$

Thus, the welfare weights determine the direction of desired redistribution (given the equal

pay condition) between workers on each job. Also, in the absence of savings taxation,

$$u'[x_{ij}] = \delta_j R u'[c_{ij}] \text{ for all } i, j.$$

The FOC for job outputs, z_i, are the same as given above.

3.2 Restricted First Best with Small Earnings-Dependent Savings Taxes

Given the observability of earnings, small linear taxes on savings (collected in the first period) could be set differently for high and low earners. This can for instance be implemented by the rules on retirement savings accounts, like the IRA and 401(k) in the US. The (local) desire to redistribute can be met by a small linear tax or subsidy on savings by workers on a given job with the revenues returned equally to them by raising net-of-tax earnings on the job.

Differentiating the Lagrangian with respect to a savings tax rate τ_i on those with earnings level y_i, evaluated at a zero tax level:

$$\frac{\partial \mathcal{L}}{\partial \tau_i} = \lambda \left(\sum_j f_{ij} (y_i - x_{ij}) \right) - \sum_j f_{ij} \eta_{ij} u'[x_{ij}] (y_i - x_{ij}).$$

The impact of a savings tax on the Lagrangian is made up of two pieces: the impact on the revenue constraint and the impact on utilities. Using the FOC with respect to y_i, multiplied by y_i, the derivative can be written as:

$$\frac{\partial \mathcal{L}}{\partial \tau_i} = \sum_j f_{ij} \left(\eta_{ij} u'[x_{ij}] - \lambda \right) x_{ij} = \sum_j f_{ij} g_{ij} x_{ij}.$$

Recall that

$$\sum_j f_{ij} g_{ij} = 0 \text{ for } i = h, l.$$

This implies that a tax on the savings by the two types on a given job increases welfare if the savings of the one type towards which redistribution is desirable saves sufficiently little compared to the other type.

The welfare weights imply the desired direction of redistribution within productivity

types and so the signs of g_{ij}. With equal incomes and different discount factors, we have

$$x_{ih} < x_{il}$$
$$c_{ih} > c_{il}$$

Thus, if first period utilities get the same weights for both types, $\eta_{il} = \eta_{ih}$, $g_{il} < 0 < g_{ih}$, implying a desire to redistribute to the high saver. In contrast, if second period utilities get the same weights for both types, $\eta_{il}\delta_l = \eta_{ih}\delta_h$, the signs are reversed, implying a desire to redistribute to the low saver. If there is no desire to redistribute for high (low) skill types we have $\eta_{hh}u'[x_{hh}] = \eta_{hl}u'[x_{hl}]$ ($\eta_{lh}u'[x_{lh}] = \eta_{ll}u'[x_{ll}]$). In general, with uniform weights for given discount factors, $\eta_{hi} = \eta_{li}$, we do not satisfy both conditions.

4 Second Best

We draw a distinction between restricted first-best analyses and second-best ones based on the absence or presence of IC constraints involving taking a job with lower productivity (the reverse having been ruled out by assumption). That is, the distinction depends on the observability of productivity. The prime issue in second-best analyses is determining which IC constraints are binding. We start with the further restriction, as above, that savings not be taxed. With no taxation of savings and equal pay for equal work, the IC constraint of not imitating the other discount rate type who is holding the same job does not bind. Similarly, if a high productivity worker were to take the low productivity job, the person imitated would be the one with the same discount factor. Imitation is a misnomer here since there need not be such a worker for a high skill worker to optimize savings while taking a low skill job given the assumed policy tools and information.

We add the critical assumption that earnings distribution issues are sufficiently important that at the second-best optimum (with IC constraints) the net marginal social value of first period consumption $g_{ij} \equiv \eta_{ij}u'[x_{ij}] - \lambda$ is negative for both of the worker types holding the high-skill job and positive for both of the types holding the low-skilled job. Without a binding IC constraint, this condition could not hold at the optimum as noted above.

Assumption 1 *The net marginal social values of first period consumption satisfy*

$$g_{hj} < 0, g_{lj} > 0, \textit{ for } j = h, l.$$

4.1 Second best with No Taxation of Savings

We assume that the Pareto-weights and population fractions are such that all high-skilled workers work at the high-skilled job and the desired level of redistribution to lower earners is sufficient that at least one IC constraint is binding.

$$
\begin{aligned}
&\text{Maximize}_{y,z} \quad \sum f_{ij}\eta_{ij}\left(w_j\left[y_i, R\right] - v\left[z_i/n_i\right]\right) \\
&\text{subject to:} \qquad E + \sum f_{ij}\left(y_i - z_i\right) \leq 0 \\
&\qquad\qquad\qquad w_h\left[y_h, R\right] - v\left[z_h/n_h\right] \geq w_h\left[y_l, R\right] - v\left[z_l/n_h\right] \\
&\qquad\qquad\qquad w_l\left[y_h, R\right] - v\left[z_h/n_h\right] \geq w_l\left[y_l, R\right] - v\left[z_l/n_h\right].
\end{aligned}
\tag{3}
$$

Forming a Lagrangian with μ_j the Lagrange multiplier for the corresponding IC constraint, and assuming that at the optimum each worker is assigned to the matching job, we have

$$
\begin{aligned}
\mathcal{L} = \sum_{i,j} f_{ij}\eta_{ij}\left(w_j\left[y_i, R\right] - v\left[z_i/n_i\right]\right) - \lambda \sum_{i,j} f_{ij}\left(y_i - z_i\right) \\
+ \sum_j \mu_j\left(w_j\left[y_h, R\right] - v\left[z_h/n_h\right] - w_j\left[y_l, R\right] + v\left[z_l/n_h\right]\right).
\end{aligned}
$$

Since the first-period consumption of type hj if switching to the low job equals the first-period consumption of type lj, the FOC with respect to earnings are

$$
\begin{aligned}
\sum_j f_{hj}\eta_{hj}u'\left[x_{hj}\right] + \sum_j \mu_j u'\left[x_{hj}\right] - \lambda \sum_j f_{hj} &= 0, \\
\sum_j f_{lj}\eta_{lj}u'\left[x_{lj}\right] - \sum_j \mu_j u'\left[x_{lj}\right] - \lambda \sum_j f_{lj} &= 0.
\end{aligned}
$$

Given the definition of the net social utility $g_{ij} \equiv \eta_{ij}u'\left[x_{ij}\right] - \lambda$, this implies

$$
\begin{aligned}
\sum_j f_{hj}g_{hj} &= -\sum_j \mu_j u'\left[x_{hj}\right] < 0, \\
\sum_j f_{lj}g_{lj} &= \sum_j \mu_j u'\left[x_{lj}\right] > 0.
\end{aligned}
$$

The population-weighted values add to a positive expression

$$
\sum_{i,j} f_{ij}g_{ij} = \sum_j \mu_j\left(u'\left[x_{lj}\right] - u'\left[x_{hj}\right]\right) > 0.
$$

That is, transfers which would be worth doing without an IC constraint are restricted, raising the social marginal utilities of consumption, on average, above the value of resources in the hands of the government. Since the IC constraints are on the high skilled types, on average more redistribution from the high earners to the low earners is desirable.

IC constraints Given the equal pay constraint, it follows that only one of the IC constraints is binding, and it is the one on the low discount factor type. To see this consider the difference in consumption utility from different incomes,

$$\Delta [y_h, y_l, \delta_j, R] \equiv w_j [y_h, R] - w_j [y_l, R].$$

This difference in consumption utility is increasing in the discount factor,

$$\frac{\partial \Delta [y_h, y_l, \delta_j, R]}{\partial \delta} = u [c_{hj}] - u [c_{lj}] > 0.$$

The difference in labor disutility does not depend on the discount factor. Thus if the IC constraint is binding on the low discount factor type, it is not binding on the high discount factor type. The low discount factor type values earnings in the first period less and is therefore more tempted to switch to the less productive job.

4.2 Second Best with Small Earnings-Dependent Taxes on Savings

As above, the sign of the welfare impact of introducing a small linear savings tax or subsidy depends on the welfare weights. Given observability of earnings, the small linear tax on savings could be different for high and low earners. The welfare impacts of introducing a tax on savings (collected in the first period) are obtained by differentiating the Lagrangian (with savings taxation included and the tax rates τ_i set at zero):

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \tau_h} &= \lambda \left(\sum_j f_{hj} (y_h - x_{hj}) \right) - \sum_j f_{hj} \eta_{hj} u' [x_{hj}] (y_h - x_{hj}) - \mu_l u' [x_{hl}] (y_h - x_{hl}), \\
\frac{\partial \mathcal{L}}{\partial \tau_l} &= \lambda \left(\sum_j f_{lj} (y_l - x_{lj}) \right) - \sum_j f_{lj} \eta_{lj} u' [x_{lj}] (y_l - x_{lj}) + \mu_l u' [x_{ll}] (y_l - x_{ll}).
\end{aligned}$$

That is, the impact on the Lagrangian is made up of three pieces: the impact on the revenue constraint, the impact on utilities, and the impact on the binding IC constraint.

The FOC for earnings are

$$\begin{aligned}\sum_j f_{hj}g_{hj} + \mu_l u'[x_{hl}] &= 0, \\ \sum_j f_{lj}g_{lj} - \mu_l u'[x_{ll}] &= 0.\end{aligned}$$

Multiplying these by the earnings level at the job, y_i, and substituting, we have

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \tau_h} &= \sum_j f_{hj}g_{hj}x_{hj} + \mu_l u'[x_{hl}]\, x_{hl}, \\ \frac{\partial \mathcal{L}}{\partial \tau_l} &= \sum_j f_{lj}g_{lj}x_{lj} - \mu_l u'[x_{ll}]\, x_{ll}.\end{aligned}$$

Substituting for $\mu_l u'[x_{il}]$ from the FOC for earnings, we find

$$\frac{\partial \mathcal{L}}{\partial \tau_h} = f_{hh}g_{hh}(x_{hh} - x_{hl}) > 0$$

and

$$\frac{\partial \mathcal{L}}{\partial \tau_l} = f_{lh}g_{lh}(x_{lh} - x_{ll}) < 0.$$

The signs follow from the assumption on the net social marginal utilities and the differences in savings behavior by types ih and il for $i = h, l$. The correlation between skill and discount plays no role in signing these expressions. The Proposition immediately follows.

Proposition 1 *At the second best optimum, assuming that all high skill workers hold high skill jobs and $g_{hj} < 0, g_{lj} > 0$, for $j = h, l$, then introduction of a small linear tax on savings that falls on high earners is welfare improving; and introduction of a small linear subsidy on savings that falls on low earners is welfare improving.*

One can increase the redistribution from high earners/high savers by taxing savings, but increasing net-of-tax earnings just enough that the high earners/low savers remain indifferent to job change and thus the binding IC constraint is unchanged. One can also increase the redistribution towards the low earners/high savers by subsidizing their savings, but decreasing net-of-tax earnings such that the low earners/low savers remain indifferent so that it does not become more attractive for the high earners/low savers to take the low job.

4.3 Second Best with Optimal Linear Earnings-Dependent Taxes on Savings

We have considered the introduction of small savings taxes on high and low earners. Part of the interest in this analysis comes from the possible link to the signs of the optimal taxes. Derivation of the FOC for the optimal linear savings taxes is straightforward; we show that it matches the signs of the small improvements given the additional condition that workers save more if the after-tax return to savings are higher.[4]

One difference in analysis is that changes in both the earnings and savings taxes have a first order effect on tax revenues through the behavioral change in savings. In first period units, the tax revenue from a linear savings tax τ_i levied on the savings of workers with discount factor δ_j and earnings y_i equals $\tau_i s_j\left[y_i, R\left(1-\tau_i\right)\right]$. For notational convenience, denote optimal savings $s_j\left[y_i, R\left(1-\tau_i\right)\right]$ by s_{ij}. (Given preference separability, there is no dependence on the effort to achieve gross earnings.) A second difference is that the relative size of the utility loss of a marginal increase in the savings tax compared to the utility gain of a marginal increase in earnings depends on the level of the savings tax. That is,

$$\frac{\partial w_{ij}}{\partial \tau_i}=-s_{ij}u'\left[c_{ij}\right]\delta R=\frac{-s_{ij}}{1-\tau_i}u'\left[x_{ij}\right]=\frac{-s_{ij}}{1-\tau_i}\frac{\partial w_{ij}}{\partial y_i}.$$

Forming a Lagrangian, and assuming that at the optimum each worker is assigned to the matching job, we now have

$$\begin{aligned}\mathcal{L}=&\sum_{i,j}f_{ij}\eta_{ij}\left(w_j\left[y_i,\left(1-\tau_i\right)R\right]-v\left[z_i/n_i\right]\right)-\lambda\sum_{i,j}f_{ij}\left\{\left(y_i-z_i\right)-\tau_i s_{ij}\left[y_i,\left(1-\tau_i\right)R\right]\right\}\\&+\mu_l\left(w_l\left[y_h,\left(1-\tau_h\right)R\right]-v\left[z_h/n_h\right]-w_l\left[y_l,\left(1-\tau_l\right)R\right]+v\left[z_l/n_h\right]\right).\end{aligned}$$

The FOC for earnings are

$$\begin{aligned}\sum_j f_{hj}\eta_{hj}u'\left[x_{hj}\right]+\mu_l u'\left[x_{hl}\right]-\sum_j\lambda f_{hj}\left(1-\tau_h\frac{\partial s_{hj}}{\partial y}\right)&=0,\\\sum_j f_{lj}\eta_{lj}u'\left[x_{lj}\right]-\mu_l u'\left[x_{ll}\right]-\sum_j\lambda f_{lj}\left(1-\tau_l\frac{\partial s_{lj}}{\partial y}\right)&=0.\end{aligned}$$

[4] Consideration of earnings-dependent nonlinear savings taxation would raise the issue of the degree of complexity that is interesting for policy purposes.

The FOC for savings tax rates are

$$\begin{aligned}
\sum_j f_{hj}\eta_{hj}u'\left[x_{hj}\right]\frac{s_{hj}}{1-\tau_h}+\mu_l u'\left[x_{hl}\right]\frac{s_{hl}}{1-\tau_h}-\sum_j \lambda f_{hj}\left\{s_{hj}+\tau_h\frac{\partial s_{hj}}{\partial \tau_h}\right\} &= 0,\\
\sum_j f_{lj}\eta_{lj}u'\left[x_{lj}\right]\frac{s_{lj}}{1-\tau_l}-\mu_l u'\left[x_{ll}\right]\frac{s_{ll}}{1-\tau_l}-\sum_j \lambda f_{lj}\left\{s_{lj}+\tau_l\frac{\partial s_{lj}}{\partial \tau_l}\right\} &= 0.
\end{aligned}$$

Denote by $R_h \equiv R(1-\tau_h)$ and $R_l \equiv R(1-\tau_l)$ the after-tax returns to savings for respectively the high and low skill types. Combining the first order conditions as before, we find that the optimal linear savings tax is such that

$$f_{hh}g_{hh}\left(x_{hh}-x_{hl}\right)=\tau_h\sum_j \lambda f_{hj}\left\{s_{hj}-\frac{\partial s_{hj}}{\partial y_h}s_{hl}+\frac{\partial s_{hj}}{\partial R_h}R_h\right\} \tag{4}$$

and

$$f_{lh}g_{lh}\left(x_{lh}-x_{ll}\right)=\tau_l\sum_j \lambda f_{lj}\left\{s_{lj}-\frac{\partial s_{lj}}{\partial y_l}s_{ll}+\frac{\partial s_{lj}}{\partial R_l}R_l\right\}. \tag{5}$$

The left-hand sides in equations (4) and (5) correspond to the welfare changes of introducing earnings-dependent taxes on the high earners and low earners respectively. Thus, if the sum of the terms in brackets on the right-hand side is positive, the optimal linear tax is positive if the introduction of a small tax is welfare-improving and vice versa. Since preferences are additive, $\frac{\partial s_{ij}}{\partial y_i}<1$, and so $s_{ij}-\frac{\partial s_{ij}}{\partial y_i}s_{il}>0$ for $i=h,l$. Hence, a sufficient condition for the right-hand side term to be positive is that savings are increasing in the after-tax return, $\frac{\partial s_{ij}}{\partial R_i}\geq 0$.

Proposition 2 *At the second best optimum, assuming that savings are increasing in the after-tax returns, all high skill workers hold high skill jobs, and* $g_{hj}<0, g_{lj}>0$ *for* $j=h,l$, *the optimal linear savings tax is positive for the high earners and negative for the low earners.*

4.4 Mechanism Design Optimum in Tenhunen and Tuomala (2008)

As noted above, Tenhunen and Tuomala (2008) consider two-, three- and four-types models with hours chosen by workers. They derive the mechanism design optimal allocations assuming CES preferences with varying correlations between discount and skill, with implicit marginal taxes shown in their Figure 1. For all but very high correlation, they find that savings are implicitly marginally taxed for the high skill worker with low discount factor (type 3), savings are implicitly subsidized for the low skill worker with high discount factor (type 2), and there are no other marginal savings distortions. With very high correlation, the low

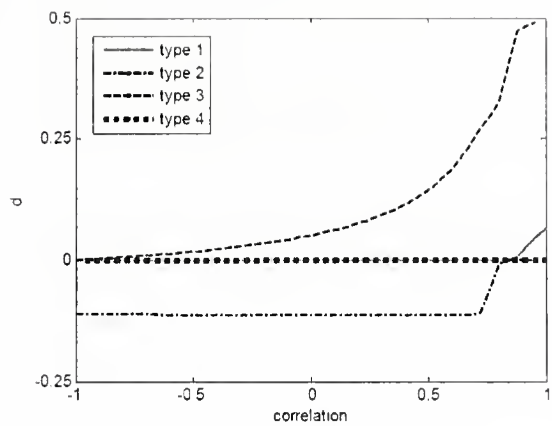

FIGURE 1: Marginal tax on savings in the welfarist case

skill worker with low discount factor (type 1) is implicitly taxed, which also happens in the two type model, which has perfect correlation. The potential relevance of the pattern we find is enhanced by their findings. In contrast with our model, the mechanism design optimal allocation allows distortion of the savings of each type separately. As long as the correlation is not too high, on average the savings tax is positive for the high skill and negative for the low skill types.

4.5 Robustness

We consider three extensions to highlight the extent of robustness of the main propositions. First, we allow different ability levels for the two high earner types. Second, we allow different discount factors for all four types. Third, we show how the analysis extends to three skill levels in workers and jobs, preserving the various assumptions.

Different Ability Levels of the High Earners In the four-types model above, we assume that the two types with high skill have exactly the same skill. As long as the high skill type with high discount factor has higher skill than the high skill type with low discount factor, Proposition 1 continues to hold. However, if the type with low discount factor is sufficiently more skilled, the type with high discount factor may be more tempted to switch to the low earner job for which less output is required. For given skill of type hh, n_{hh}, this reversal of which IC constraint is binding holds when the ability level n_{hl} of type hl is higher than $\hat{n}_{hl}$ ($> n_{hh}$), where the cut-off level $\hat{n}_{hl}$ is such that the IC constraint is just

binding on both types,

$$\{w_l [y_h, R] - v (z_h/\hat{n}_{hl})\} - \{w_l [y_l, R] - v (z_l/\hat{n}_{hl})\} = \{w_h [y_h, R] - v (z_h/n_{hh})\} - \{w_h [y_l, R] - v (z_l/n_{hh})\}.$$

With $v [z/n]$ convex, the difference in labor disutilities between jobs, $\{v (z_l/n) - v (z_h/n)\}$, is decreasing in n. Hence, for values of n_{hl} higher than $\hat{n}_{hl}$ the IC constraint is more stringent for the high discount saver. In this case, a savings subsidy on the high earners and a savings tax on the low earners are welfare improving. This is the opposite of Proposition 1.

Different Discount Factors among the High and Low Savers With job-specific earnings and no taxation of savings, a high skill worker considering switching to the low job chooses optimal savings without needing to match any particular worker holding the low job. Thus, with the same skill among high earners, the gain from switching to the low job is always higher for the high skill worker who has lower preference for savings, regardless of the discount rates among the low skill workers. We continue to have a welfare gain from introducing taxation of savings among high earners as in Proposition 1.

Subsidization of savings of low earners will continue to generate a welfare gain as long as the discount factor of the high-skill low-saver is small enough relative to the distribution of discount factors among holders of the low skill job. Denoting by $\widetilde{x}_{hl}$ the first-period consumption of the high-skilled low saver if taking the low skill job, the FOC for earnings on that job is:

$$\sum_j f_{lj} g_{lj} - \mu_l u' [\widetilde{x}_{hl}] = 0.$$

The impact of a savings tax on low earners is

$$\frac{\partial \mathcal{L}}{\partial \tau_l} = \sum_j f_{lj} g_{lj} x_{lj} - \mu_l u' [\widetilde{x}_{hl}] \widetilde{x}_{hl}.$$

Comparing the consumption in the IC constraint with a weighted average of consumptions among low earners, this derivative is negative (and the gain from the subsidization of the savings of low earners in Proposition 1 continues to hold) if and only if $\widetilde{x}_{hl} > \overline{x}_l$, where

$$\overline{x}_l = \frac{\sum_j f_{lj} g_{lj} x_{lj}}{\sum_j f_{lj} g_{lj}}.$$

With the net marginal social values assumption, $g_{lj} > 0$, for $j = h, l$, $\overline{x}_l$ is a proper weighted

average of the x_{lj}. Since the discount rates for the marginal high skill type may well be too high to meet this condition, we consider the tax of the savings of higher earners to be a more robust policy conclusion than the subsidization of the savings of low earners.

We are exploring two extensions to the basic model, one with an education choice and one with a continuum of worker types. In both cases, preliminary work suggests the same pattern of greater robustness of the taxation of higher earners than of the subsidization of lower earners.[5]

Three Ability Levels, Three Jobs We introduce an intermediate skill level in the model. We extend the assumption that welfare weights and population fractions are such that at the optimum all the high skilled are on the most productive job to also have all the intermediate skilled on the intermediate job. We again consider the case in which agents may be tempted to switch to jobs designed for less skilled people. Only two downward constraints are relevant though.

First, as above, for two agents with the same skill, but different discount factors, the IC constraint is slack for the type with the higher discount factor if it is binding for the type with the lower discount factor. The reason is that, with

$$\Delta\left[y_1, y_2, \delta_j, R\right] \equiv w_j\left[y_1, R\right] - w_j\left[y_2, R\right],$$

$$\text{both } \frac{\partial \Delta\left[y_m, y_l, \delta_j, R\right]}{\partial \delta} > 0 \text{ and } \frac{\partial \Delta\left[y_h, y_i, \delta_j, R\right]}{\partial \delta} > 0 \text{ for } i = m, l.$$

Second, with $v\left[z/n\right]$ convex, we have a similar condition for the difference in the disutility of labor between jobs. That is, with

$$\Delta'\left[z_h, z_l, n\right] \equiv v\left[z_h/n\right] - v\left[z_l/n\right],$$

$$\frac{\partial \Delta'\left[z_h, z_l, n\right]}{\partial n} = \left(-v'\left[z_h/n\right] z_h + v'\left[z_l/n\right] z_l\right)/n^2 < 0.$$

Thus, for two agents with the same discount factor, the IC constraint of switching to the low-skilled job is slack for the type with the highest ability if it is satisfied for the type with the intermediate ability switching to the low-skilled job and for the type with highest ability switching to the intermediate job. That is, the local IC constraints imply the global IC constraint.

In a similar way as for the four-types model, we can set up the Lagrangian for the

[5] With heterogeneity in discount factors, people who discount the future less may choose to invest more in education. If only education determines a worker's skill level, high-skilled workers have higher discount factors than low-skilled workers.

constrained maximization problem. The two relevant IC constraints are

$$\begin{aligned} w_l[y_h, R] - v[z_h/n_h] &\geq w_l[y_m, R] - v[z_m/n_h], \\ w_l[y_m, R] - v[z_m/n_m] &\geq w_l[y_l, R] - v[z_l/n_m]. \end{aligned}$$

The impact of the introduction of earnings-dependent savings taxes on the Lagrangian equals respectively

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \tau_h} &= f_{hh} g_{hh} (x_{hh} - x_{hl}) > 0, \\ \frac{\partial \mathcal{L}}{\partial \tau_m} &= f_{mh} g_{mh} (x_{mh} - x_{ml}) \gtrless 0, \\ \frac{\partial \mathcal{L}}{\partial \tau_l} &= f_{lh} g_{lh} (x_{lh} - x_{ll}) < 0. \end{aligned}$$

This implies that Proposition 1 continues to hold for the high earners and the low earners. The following Proposition applies for the intermediate earners.

Proposition 3 *In a model with three ability levels and three jobs, the introduction of a small linear tax (subsidy) on savings that falls on the intermediate earners is welfare improving if redistribution from (to) the intermediate earners to (from) general revenues is welfare improving.*

Proposition 3 implies that there is a single sign change in the response of welfare to taxing savings as a function of earnings. This result generalizes for more than three jobs as well, if the welfare weights are non-increasing in skill. The savings of workers with earnings below a given level are subsidized, the savings of workers with earning above that level are taxed. The result depends on the assumption that types with the same skill are at the same job, which becomes increasingly strained with many jobs.

The single sign change of the welfare impact of introducing a savings tax as a function of earnings also holds for the optimal linear earnings-dependent savings taxes when workers have CRRA preferences, $u[x] = \frac{x^{1-\gamma}}{1-\gamma}$, and $\gamma < 1$. With logarithmic preferences, $u[x] = \log[x]$, the optimal savings tax rate is strictly increasing in the earnings of workers if they are uniformly distributed across jobs, $f_{ij} = f_j$ for $\forall i, j$. Since for logarithmic preferences $s_{ij} = \frac{\partial s_{ij}}{\partial y_i} y_i$ and $\frac{\partial s_{ij}}{\partial R_i} = 0$, the optimal tax on the savings of earners at job i satisfies

$$f_{ih} g_{ih} (x_{ih} - x_{il}) = \tau_i \sum_j \lambda f_{ij} \frac{\partial s_{ij}}{\partial y_i} x_{il}.$$

With $f_{ij} = f_j$ for $\forall i, j$ and first-period consumption $x_{ij} = \frac{1}{1+\delta_j} y_i$, we find the following

expression for the optimal savings tax,

$$\tau_i = \frac{f_h(\delta_l - \delta_h)}{\sum_j \lambda f_j \frac{\delta_j}{1+\delta_j}} \left(\frac{\eta_{ih}}{y_i} - \frac{\lambda}{1+\delta_h} \right).$$

Since $\delta_h > \delta_l$, with the welfare weights non-increasing in skill, this implies that the optimal linear savings tax is increasing in earnings,

$$\frac{\partial \tau_i}{\partial y_i} > 0.$$

5 Second Best with Uniform Taxes on Savings

Proposition 1 leaves the natural question of what to do with a Nordic dual income tax where the tax rate on savings is required to be the same for both earnings levels. Adding the responses to the two separate tax changes, we have

$$\frac{\partial \mathcal{L}}{\partial \tau} = \frac{\partial \mathcal{L}}{\partial \tau_h} + \frac{\partial \mathcal{L}}{\partial \tau_l} = f_{hh} g_{hh} (x_{hh} - x_{hl}) + f_{lh} g_{lh} (x_{lh} - x_{ll}).$$

In contrast with the earnings-varying tax on savings, the correlation between skill and discount factor plays a role here.

If there is no desire to redistribute within a job, $g_{ih} = g_{il}$, for $i = h, l$, then

$$\begin{aligned} f_{hh} g_{hh} &= \frac{f_{hh}}{\sum_j f_{hj}} \sum_j f_{hj} g_{hj} = -\frac{f_{hh}}{\sum_j f_{hj}} \mu_l u' [x_{hl}], \\ f_{lh} g_{lh} &= \frac{f_{lh}}{\sum_j f_{lj}} \sum_j f_{lj} g_{lj} = \frac{f_{lh}}{\sum_j f_{lj}} \mu_l u' [x_{ll}]. \end{aligned}$$

Thus, the welfare impact of a change in the uniform tax on savings equals

$$\frac{\partial \mathcal{L}}{\partial \tau} = \mu_l \left(\frac{f_{hh}}{\sum_j f_{hj}} u' [x_{hl}] (x_{hl} - x_{hh}) - \frac{f_{lh}}{\sum_j f_{lj}} u' [x_{ll}] (x_{ll} - x_{lh}) \right).$$

It is convenient to write this as

$$\frac{\partial \mathcal{L}}{\partial \tau} = \mu_l \frac{f_{lh}}{\sum_j f_{lj}} u' [x_{ll}] (x_{ll} - x_{lh}) \left(\frac{f_{hh} / \sum_j f_{hj}}{f_{lh} / \sum_j f_{lj}} \Omega - 1 \right),$$

with

$$\Omega \equiv \frac{u' [x_{hl}] (x_{hl} - x_{hh})}{u' [x_{ll}] (x_{ll} - x_{lh})} > 0.$$

Since $x_{ll} > x_{lh}$,

$$sign\left(\frac{\partial \mathcal{L}}{\partial \tau}\right) = sign\left(\frac{f_{hh}/\sum_j f_{hj}}{f_{lh}/\sum_j f_{lj}}\Omega - 1\right).$$

The sign of this expression depends on the distribution of types and the ratio of the weights, Ω. That is, $\Omega \geq 1$ is a sufficient condition for a positive correlation between aspects, $\frac{f_{hh}}{\sum_j f_{hj}} > \frac{f_{lh}}{\sum_j f_{lj}}$, to imply that introducing a savings tax increases social welfare.

Assuming homothetic preferences, so that $\frac{x_{hl}}{x_{ll}} = \frac{x_{hh}}{x_{lh}}$, the expression for Ω becomes $\frac{u'[x_{hl}]x_{hl}}{u'[x_{ll}]x_{ll}}$. This expression is equal to one for the log utility function. For CRRA preferences, $u[x] = \frac{x^{1-\gamma}}{1-\gamma}$, we find

$$\Omega = \left(\frac{x_{hl}}{x_{ll}}\right)^{1-\gamma}.$$

Thus, if the relative risk aversion γ is smaller than 1, then $\Omega \geq 1$ and a positive correlation between ability and discount factor (i.e. $\frac{f_{hh}}{\sum_j f_{hj}} > \frac{f_{lh}}{\sum_j f_{lj}}$) implies that $\frac{\partial \mathcal{L}}{\partial \tau}$ is positive. If γ is larger than 1, the sign of $\frac{\partial \mathcal{L}}{\partial \tau}$ depends on the size of the correlation and the magnitude of the earnings difference between jobs. Conversely, when the correlation is negative, $\frac{\partial \mathcal{L}}{\partial \tau}$ is negative if γ is greater than 1.[6] This implies the following proposition.

Proposition 4 *If there is no desire to redistribute within a job, $g_{ih} = g_{il}$, for $i = h, l$, with CRRA preferences, a uniform small tax on savings increases welfare if the relative risk aversion is smaller than one and the correlation between ability and discount factor is positive. A uniform small subsidy on savings increases welfare if the relative risk aversion is greater than one and the correlation between ability and discount factor is negative.*

Corollary 1 *If there is no desire to redistribute within a job, $g_{ih} = g_{il}$, for $i = h, l$, with logarithmic preferences, a uniform small tax (subsidy) on savings increases welfare if and only if the correlation between ability and discount factor is positive (negative).*

As with the earnings-varying taxes, the sign result for introducing a uniform tax matches that for optimal linear taxation in some interesting cases. Denote by $R_\tau \equiv R(1-\tau)$ the after-tax returns to savings and by s_{ij} the savings of type ij as a function of after-tax earnings and the after-tax interest rate. Setting the derivative of the Lagrangian with respect to τ equal to zero, we find the following condition for the optimal linear tax,

$$f_{hh}g_{hh}(x_{hh} - x_{hl}) + f_{lh}g_{lh}(x_{lh} - x_{ll}) = \tau \sum_{i,j} \lambda f_{ij} \left\{ s_{ij} - \frac{\partial s_{ij}}{\partial y_i} s_{il} + \frac{\partial s_{ij}}{\partial R_\tau} R_\tau \right\}.$$

[6] For CARA preferences, $\frac{\partial \mathcal{L}}{\partial \tau}$ is negative when the correlation between ability and discount factor is negative. When the correlation is positive, $\frac{\partial \mathcal{L}}{\partial \tau}$ is positive if the absolute risk aversion is sufficiently small.

If the sum of the terms in brackets is positive, we have that the optimal uniform tax is positive if the introduction of a small uniform tax is welfare improving. This is the case for logarithmic preferences and CRRA preferences with relative risk aversion $\gamma < 1$.

Proposition 5 *If there is no desire to redistribute within a job, $g_{ih} = g_{il}$, for $i = h, l$, with logarithmic preferences or CRRA preferences with $\gamma < 1$, the optimal linear uniform tax on savings is positive if the correlation between ability and discount factor is positive.*

6 Two Types

Using a model with two types of workers, a high-skilled worker with high discount factor and a low-skilled worker with low discount factor, the empirical finding of a positive correlation between skill and savings rates is treated as a perfect correlation. In this model, if there is positive (negative) marginal earnings taxation then there is a gain from introducing positive (negative) marginal savings taxation. The corollary is that introducing savings taxation is a gain if redistribution goes from the high earner to the low earner. The full mechanism design optimum has the same property. The source of this inference does not seem robust to realistic diversity in the economy. With two-dimensional heterogeneity, there are low earners with both high and low savings rates. If a high earner can imitate the savings of a low earner with the same savings propensities, a savings tax on the low earner does not work to discourage the high earner from imitating. Thus to model less-than-perfect correlation, we use the four-types model with high and low earners with both high and low concern for the future. We report the results for two types here to mark the contrast with the four types model. The proof parallels that of the same result for the mechanism design optimum, which is in Diamond (2003). We consider the second-best Pareto frontier with the types referred to as 1 and 2.

Proposition 6 *In a two-types model without taxes on savings and with $sign\,(\delta_1 - \delta_2) = sign\,(n_1 - n_2)$, the introduction of a small linear tax (subsidy) on savings at a given earnings level is welfare improving if and only if earnings at that level are marginally taxed (subsidized).*

Corollary 2 *In a two-types model without taxes on savings and with $sign\,(\delta_1 - \delta_2) = sign\,(n_1 - n_2)$, the introduction of a savings tax on the lower earner is welfare improving if redistribution goes from the higher earner to the lower earner.*

The proposition combines the properties of the mechanism design optima in the two separate two-types models with heterogeneity in one dimension. When both types have the

same discount factor, but different abilities, the earnings of the potentially imitated type are marginally taxed or subsidized if the ability of that type is lower or higher respectively (Mirrlees, 1971). Similarly, when both types have the same ability, but different discount factors, the savings of the potentially imitated type are marginally taxed or subsidized if the discount factor for that type is lower or higher respectively. If both types have the same discount factor, distorting savings does not help separate the two types (Atkinson and Stiglitz, 1976). Similarly, if both types have the same ability, earnings are not subject to marginal taxation.[7] However, when the two types differ in both ability and discount factor, both the marginal taxation (or subsidization) of earnings and the marginal taxation (or subsidization) of savings is used to separate types.

7 Preferences and IC Constraints

Above we used the utility functions $u[x] + \delta_j u[c] - v[z/n_j]$. This family of utility functions has the property that those with higher savings rates (larger values of δ_j) are more willing to increase work for a given amount of additional pay. But that is not the only way in which the savings and labor supply decisions can be connected. For example, with the utility functions $(u[x] + \delta_j u[c]) / \delta_j - v[z/n_j] = u[x] / \delta_j + u[c] - v[z/n_j]$, the relationship is reversed - those with higher savings rates are less willing to increase work for additional pay. If we had assumed this class of functions, then we would have reversed the pattern of desirable savings taxes in Proposition 1 - having the IC constraint bind for the high saver would imply that it is not binding for the low saver, implying, in turn, that there should be a subsidy of savings for high earners and a tax on savings for low earners. More generally, a one-dimensional family of separable utility functions, $U[\phi[x, c, j], z, j]$, can have any pattern between the variation in the subutility function of consumption and the variation in the interaction between consumption and labor. This raises the question of identifying an empirical basis for distinguishing which case is more relevant. That it is standard practice to write utility in the form employed does not, by itself, shed light on its empirical reality.

While the formal model has consumption and work simultaneous in the first period, experience in real time is different. Generally, work precedes pay, which precedes spending it (but not borrowing against it). So modeling in continuous time would naturally have a similar role for discounting on both aspects - saving and willingness to work. But that does not rule out the possibility that the preferences of high and low savers differ in other ways than just

[7] This can be considered an implication of the Atkinson-Stiglitz theorem, since there is separability and everyone has the same subutility function over first period earnings and consumption. Notice that if a savings tax is not allowed, the two types can be usefully separated by an earnings tax. The marginal tax on the earnings of the potentially imitated type is positive if and only if that type saves less for the same earnings.

a discount rate on otherwise identical utility and disutility functions (assuming additivity). There are reasons based on casual empiricism for supporting the appropriateness of using the formulation employed above. Modeling savings with rationality and discounting combines underlying preferences and issues of self-control. As discussed in Banks and Diamond (2008), psychological analyses suggest these are mixed together. We see no reason to think that this does not apply to working as well as to consuming - whether that is working for later consumption or working to influence future work opportunities. That is, working (at a job with disutility) involves self-control for a future payoff. And saving involves self-control. So those with less difficulty in self-control may show greater willingness to both work and save, which would be captured in the standard utility function expression. In a richer model, human capital investment involves discounting in a similar way to savings decisions and so may generate the pattern in the standard model structure, although formal modeling would distinguish between human and financial capital accumulations.

It is not easy to find data applying directly to this issue. The question we want to answer is whether, for a given level of skill, those with higher savings rates tend to have greater labor supply functions. A complication in looking at data comes from the differences in circumstances with age, which we address by considering separate age cells. We report a few correlations supportive of a positive correlation among savings propensities, discounting and earnings abilities using the Survey of Consumer Finances, which includes some questions on time horizon and savings practice.[8] We also report some correlations with work effort. Before turning to the data, we briefly consider a three-period version of the two-period model we have been considering. This will bring out some of the complications in interpreting the data at different ages. There is also a complication in interpreting the data across education levels. Education choices reflect both ex ante "skill" and discount rate and then affect wage rates, which matter for later taxation. Presumably, the level of completed education is increasing in both ex ante skill and discount factor, on average. In addition to affecting ex post skill, education may affect one's discount rate thereafter. Thus education is a proxy for both skill and discount rate and can not be used in a simple way to distinguish between them. A further difficulty in interpreting the correlations is that education is a discrete variable while skill is continuous and varying within education classes.

7.1 Three-period Model

We set up a three-period model with the same preference structure as the two-period model analyzed, assuming the same discounting for the utility of consumption and the disutility of

[8] For discussion of correlations with experimentally measured discount rates, see Chabris et al, 2008.

work and allowing for different skills in the two working periods.

$$u[x_1] + \delta u[x_2] + \delta^2 u[c] - v[z_1/n_1] - \delta v[z_2/n_2]$$

Considering the special case with

$$u[x] = \log[x] \text{ and } v[z/n] = k(z/n)^{\beta+1}/(\beta+1),$$

we have the time series of consumption and earnings behavior, assuming that hours are a control variable and the marginal tax rate is constant over time (with derivation in the Appendix):

$$\begin{aligned}
\frac{x_2}{x_1} &= \frac{c}{x_2} = \delta R \\
\frac{z_2}{z_1} &= (\delta R)^{-1/\beta}\left(\frac{n_2}{n_1}\right)^{1+1/\beta} \\
\frac{z_2/n_2}{z_1/n_1} &= (\delta R)^{-1/\beta}\left(\frac{n_2}{n_1}\right)^{1/\beta}.
\end{aligned}$$

Thus, those with higher discount factors have more rapidly growing consumption but less rapidly growing earnings (for given skills). This suggests that the cross section pattern of earnings and work effort may be different at different ages.

The cross-section pattern of time series behavior may be more illuminating than that of single-period behavior since the single-period cross section patterns are dependent on the full pattern over time in skills. If we added uncertain rates of return to the model, we would also be concerned about income effects in both consumption and earnings choices. Consideration of wealth or wealth/earnings ratios are also affected by the time series pattern of skills. But we do not explore these issues, just reporting simple correlations.

7.2 Data Analysis

We first consider the relations among discounting, saving, education and age. We use the SCF panels in 1998, 2001 and 2004, containing information on 13,266 households in total. For savings rates we consider two proxies. The first proxy is the logarithm of the ratio of net worth to earnings for households. The second proxy is whether people report that they save regularly or not.[9] The sample is divided into age-education cells (5-year age groupings from

[9]Subjects can choose among different statements. We use for this second proxy whether subjects confirm the statement: "Save regularly by putting money aside each month." The results are similar (with sign reversal) with the statement "Don't save - usually spend about as much as income."